



Aalborg Universitet

AALBORG UNIVERSITY  
DENMARK

## Pilot Decontamination Through Pilot Sequence Hopping in Massive MIMO Systems

Sørensen, Jesper Hemming; De Carvalho, Elisabeth

*Published in:*  
IEEE Global Telecommunications Conference (GLOBECOM)

*DOI (link to publication from Publisher):*  
[10.1109/GLOCOM.2014.7037313](https://doi.org/10.1109/GLOCOM.2014.7037313)

*Publication date:*  
2014

*Document Version*  
Accepted author manuscript, peer reviewed version

[Link to publication from Aalborg University](#)

*Citation for published version (APA):*  
Sørensen, J. H., & De Carvalho, E. (2014). Pilot Decontamination Through Pilot Sequence Hopping in Massive MIMO Systems. In *IEEE Global Telecommunications Conference (GLOBECOM)* (pp. 3285 - 3290). IEEE. GLOBECOM - conference record / IEEE Global Telecommunications Conference  
<https://doi.org/10.1109/GLOCOM.2014.7037313>

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

### Take down policy

If you believe that this document breaches copyright please contact us at [vbn@aub.aau.dk](mailto:vbn@aub.aau.dk) providing details, and we will remove access to the work immediately and investigate your claim.

# Pilot Decontamination Through Pilot Sequence Hopping in Massive MIMO Systems

Jesper H. Sørensen and Elisabeth de Carvalho

Aalborg University, Department of Electronic Systems, Fredrik Bajers Vej 7, 9220 Aalborg, Denmark

E-mail: {jhs,edc}@es.aau.dk

**Abstract**—This work concerns wireless cellular networks applying massive multiple-input multiple-output (MIMO) technology. In such a system, the base station in a given cell is equipped with a very large number (hundreds or even thousands) of antennas and serves multiple users. Estimation of the channel from the base station to each user is performed at the base station using an uplink pilot sequence. Such a channel estimation procedure suffers from pilot contamination. Orthogonal pilot sequences are used in a given cell but, due to the shortage of orthogonal sequences, the same pilot sequences must be reused in neighboring cells, causing pilot contamination. The solution presented in this paper suppresses pilot contamination, without the need for coordination among cells. Pilot sequence hopping is performed at each transmission slot, which provides a randomization of the pilot contamination. Using a modified Kalman filter, it is shown that such randomized contamination can be significantly suppressed. Comparisons with conventional estimation methods show that the mean squared error can be lowered as much as an order of magnitude at low mobility.

## I. INTRODUCTION

Multiple-input multiple-output (MIMO) technology [1] is finding its way into practical systems, like LTE and its successor LTE-Advanced. It is a key component for these systems' ability to improve the spectral efficiency. The success of MIMO technology has motivated research in extending the idea of MIMO to cases with hundreds, or even thousands of antennas, at transmitting and/or receiving side. This is often termed *massive MIMO*. In mobile communication systems, like LTE, the more realistic scenario is to have a massive amount of antennas only at the base station (BS), due to the physical limitations at the user equipment (UE). It has been shown that such a system [2], in theory, can eliminate entirely the effect of small-scale fading and thermal noise, when the number of BS antennas goes to infinity. The only remaining impairment is inter-cell interference, caused by imperfect channel state information (CSI), which is a result of non-orthogonality of training pilots used to gather the CSI. This is often referred to as *pilot contamination*. It is considered as one of the major challenges in massive MIMO systems [3].

Mitigation of pilot contamination has been the focus of several works recently. These fall into two categories; one with coordination among cells and one without. The first category includes [4], where it is utilized that the desired and interfering signals can be distinguished in the channel covariance matrices, as long as the angle-of-arrival spreads of desired and

interfering signals do not overlap. A pilot coordination scheme is proposed to help satisfying this condition. The work in [5] utilizes coordination among base stations to share downlink messages. Each BS then performs linear combinations of messages intended for users applying the same pilot sequence. This is shown to eliminate interference when the number of base station antennas goes to infinity.

The category without coordination also includes notable contributions. A multi-cell precoding technique is used in [6] with the objective of not only minimizing the mean squared error of the signals of interest within the cell, but also minimizing the interference imposed to other cells. In [7] it is shown that channel estimates can be found as eigenvectors of the covariance matrix of the received signal when the number of base station antennas grows large and the system has "favorable propagation". The work in [8–11] is based on examining the eigenvalue distribution of the received signal to identify an interference free subspace on which the signal is projected. It is shown that an interference free subspace can be identified when certain conditions are fulfilled concerning the number of base station antennas, user equipment antennas, channel coherence time and the signal-to-interference ratio.

The major contribution of this paper is a pilot decontamination, which does not require inter-cell coordination, and is able to exploit past pilot signals. It is based on pilot sequence hopping performed within each cell. Pilot sequence hopping means that every user chooses a new pilot sequence in each transmission slot. Consider a user of interest and the effect of the inter-cell pilot contamination when pilot sequence hopping is applied. At each transmission slot, the pilot signal of the user is contaminated by a different set of interfering users. Hence channel estimation at each transmission slot is affected by a different set of interfering channels. If channel estimation is carried out based solely on the pilot sequence of the current slot, then pilot sequence hopping does not bring any gain. The key in our solution is a channel estimation that incorporates multiple time slots so that it can benefit from randomization of the pilot contamination. Recent work utilizing temporal correlation for channel estimation is found in [12], although not in combination with pilot hopping and not with the purpose of mitigating pilot contamination.

Consider the simple example, where the channel of the UE of interest is time-invariant. Its estimation is performed across multiple time slots. Specifically, the resulting channel estimate is the average of the estimates across the time slots. In the

averaging process, the contamination signal is averaged out. Note that, if the contamination signal remains constant across the time slots, i.e. there is no hopping, this averaging brings no benefit (except an averaging of the receive noise).

When the channel is time-variant and correlated across time slots, it remains possible to exploit the information about the channel across time slots by an appropriate filtering and benefit from contamination randomization. In this paper, channel estimation across multiple time slots is performed using a modified version of the Kalman filter, which is capable of tracking the channel and the channel correlation. The level of contamination suppression depends on the channel correlation between slots of the UE of interest as well as the contaminants. In LTE, channel correlation between time slots is large even at medium-high speeds, making the proposed solution very efficient.

The remainder of this paper is organized as follows. Section II presents the applied system model and the problem of pilot contamination. The proposed solution is described in section III and evaluated and compared to existing solutions in section IV. Finally, conclusions are drawn in section V.

## II. SYSTEM MODEL

In this work we denote scalars in lower case, vectors in bold lower case and matrices in bold upper case. A superscript “ $T$ ” denotes the transpose and a superscript “ $H$ ” denotes the conjugate transpose.

This work treats a cellular system consisting of  $L$  cells with  $K$  users in each cell. A massive MIMO scenario is considered, where the BS has  $M$  antennas and the UE has a single antenna. We restrict our attention to the channel estimation performed in a single cell, which we term “the cell of interest” and assign the index “0”. The channel between the BS in the cell of interest and the  $k$ ’th user in the  $\ell$ ’th cell is denoted  $\mathbf{h}^{k\ell} = [h^{k\ell}(1) h^{k\ell}(2) \dots h^{k\ell}(M)]$ , where the individual channel coefficients are complex scalars. Note that for  $\ell > 0$ ,  $\mathbf{h}^{k\ell}$  refers to a channel between the BS of interest and a UE connected to a different base station. We furthermore restrict our attention to the estimation of a single channel coefficient, hence a channel is denoted as the complex scalar  $h^{k\ell}$ . The work easily extends to vector estimations, in which case spatial correlation can be exploited for improved performance. A rich scattering environment is assumed, such that  $h^{k\ell}$  can be modeled using Clarke’s model [13], hence

$$h^{k\ell} = \frac{1}{\sqrt{N_s}} \sum_{m=1}^{N_s} e^{j2\pi f_d t \cos \alpha_m + \phi_m}, \quad (1)$$

where  $N_s$  is the number of scatterers,  $f_d$  is the maximum Doppler shift,  $\alpha_m$  and  $\phi_m$  is the angle of arrival and initial phase, respectively, of the wave from the  $m$ ’th scatterer. Both  $\alpha_m$  and  $\phi_m$  are i.i.d. in the interval  $[-\pi, \pi)$  and  $f_d = \frac{v}{c} f_c$ , where  $v$  is the speed of the UE,  $c$  is the speed of light and  $f_c$  is the carrier frequency.

In a massive MIMO system, collection of channel state information (CSI) is performed using uplink pilot training.

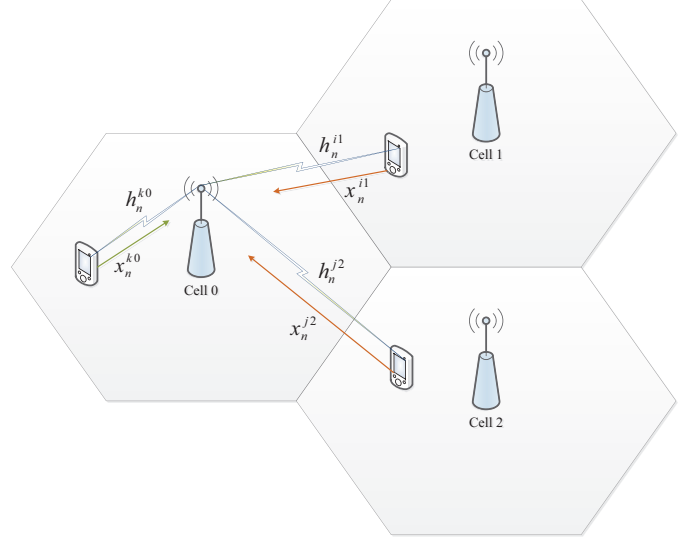


Fig. 1. A cellular system with three cells. Cell 0 is of interest and the neighboring cells will potentially cause interference (red arrows).

The CSI achieved this way is utilized in both downlink and uplink transmissions based on the channel reciprocity assumption. We define a pilot training period followed by an uplink and a downlink transmission period as a time slot. See Fig. 2 for an example of a transmission schedule with two time slots. During the  $n$ ’th pilot training period, the  $k$ ’th user in the  $\ell$ ’th cell transmits a pilot sequence  $\mathbf{x}_n^{k\ell} = [x_n^{k\ell}(1) x_n^{k\ell}(2) \dots x_n^{k\ell}(\tau)]^T$ , where  $\tau$  is the pilot sequence length. Ideally, all pilot sequences in the entire system are orthogonal, in order to avoid interference. However, this would require pilot sequences of at least length  $L \cdot K$ , which in most practical systems is not feasible. Instead, orthogonality within each cell only is ensured, i.e.  $\tau = K$ , thereby dealing with the potentially strongest sources of interference. As a result, all cells use the same set of pilots, potentially causing interference from neighboring cells. This is referred to as pilot contamination. We define the contaminating set,  $\mathcal{C}_n^{k\ell}$ , as the set of all pairs  $i, j$ , which identify all UEs applying the same pilot sequence in the  $n$ ’th time slot as the  $k$ ’th user in the  $\ell$ ’th cell. Hence,  $\mathbf{x}_n^{ij} = \mathbf{x}_n^{k\ell} \forall i, j \in \mathcal{C}_n^{k\ell}$ .

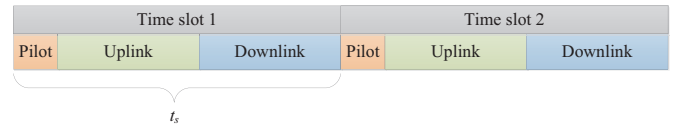


Fig. 2. Scheduling example.

The pilot signal received by the BS of interest, concerning the  $k$ ’th user in the  $n$ ’th time slot can be expressed as

$$\mathbf{y}_n^{k0} = h_n^{k0} \mathbf{x}_n^{k0} + \sum_{i,j \in \mathcal{C}_n^{k0}} h_n^{ij} \mathbf{x}_n^{ij} + \mathbf{z}_n^{k0}, \quad (2)$$

where  $\mathbf{z}_n^{k0} = [z_n^{k0}(1) z_n^{k0}(2) \dots z_n^{k0}(\tau)]^T$  and  $z_n^{k0}(j)$  are

circularly symmetric Gaussian random variables with zero mean and unit variance for all  $j$ . Here, only signals leading to contamination are included in the sum term, since any  $h_n^{ij} \mathbf{x}_n^{ij} \forall i, j \notin \mathcal{C}_n^{\ell}$  are removed when correlating with the applied pilot sequence. Hence, all contributions from the sum term are undesirable and will contaminate the CSI. Without loss of generality, we focus on the channel estimation for a single user in a single cell. Hence, in the remainder of the paper, we omit the superscript  $k$  for ease of notation.

### III. PILOT DECONTAMINATION

The solution to pilot contamination proposed in this work consists of two components:

- 1) **Pilot sequence hopping:** This component refers to random shuffling of the pilots applied within a cell. This shuffle occurs between every time slot. The purpose of this component is to *decorrelate* the contaminating signals. When pilots are shuffled, the set of contaminating users will be replaced by a new set, whose channel coefficients are uncorrelated with those of the previous set.
- 2) **Kalman filtering:** The autocorrelation of the channel coefficient of the user of interest is high at low mobility. This means that information about the value of the current channel coefficient exists not only in the most recent pilot signal, but also in past pilot signals. This can be extracted using a filter. For this purpose a Kalman filter is desirable due to its recursive structure, which provides low complexity, yet optimal performance. Additionally, since the contaminating signals have been decorrelated, the Kalman filter will suppress the impact of these signals, leading to pilot *decontamination*.

#### A. Pilot Sequence Hopping

Pilot sequence hopping is a technique where the UEs randomly switch to a new pilot sequence in between time slots. This must be coordinated with the BS, which in practice can be realized by letting the BS send a seed for a pseudorandom number generator to each UE. Random pilot sequence hopping is illustrated in Fig. 3 in the case of  $\tau = K = 5$ . Note how the identity of the contaminator changes between time slots, as opposed to a fixed pilot sequence schedule, where the contaminator remains the same UE. Consequently, the undesirable part of the pilot signal, i.e. the sum term in (2), varies rapidly between time slots compared to the variation caused by the mobility of a single contaminator in a fixed schedule. In fact, the impact of pilot sequence hopping, from a contamination perspective, can be viewed as a dramatic increase of the mobility of the contaminator. This in turn leads to a lowered autocorrelation, or decorrelation, in the contaminating signal, which is the motivation behind performing pilot sequence hopping.

The level of decorrelation is related to the time between two instances, where the same user acts as a contaminator. We refer to this as the collision distance, and we denote it  $t_c$ , see Fig. 3. Note that in the case of a fixed pilot

schedule,  $t_c = 1$ . The goal of pilot sequence hopping is to maximize  $t_c$ , either in an expected sense or maxmin sense, i.e. maximization of the minimum value. The latter can be pursued through a minimal level of coordination of pilot sequence schedules among neighboring cells. However, this work is strictly restricted to a framework with no inter-cell coordination, hence, we focus on the expected value of  $t_c$ . If pilot sequence hopping is performed at random and  $\tau = K$ , then  $t_c$  follows a geometric distribution, such that

$$P(t_c = d) = (1 - p)^{d-1}p, \quad d = 1, 2, \dots, \\ p = \frac{1}{K}, \quad (3)$$

where  $P(t_c = d)$  is the probability that the collision distance is  $d$  and  $p$  is the probability of a given UE being the next contaminator. The expected value of  $t_c$ ,  $\mathbb{E}[t_c]$ , is then found as

$$\mathbb{E}[t_c] = \sum_{d=1}^{\infty} d(1 - p)^{d-1}p \\ = \sum_{d=1}^{\infty} d \left( \frac{K-1}{K} \right)^{d-1} \frac{1}{K} \\ = K. \quad (4)$$

Hence, the expected collision distance increases with the number of users/pilots per cell, which follows intuition.

		Cell 0	Cell 1				
		UE of interest	UE 1	UE 2	UE 3	UE 4	UE 5
Time slot	1	$X_3$	$X_3$	$X_5$	$X_1$	$X_2$	$X_4$
	2	$X_2$	$X_1$	$X_4$	$X_5$	$X_3$	$X_2$
	3	$X_5$	$X_1$	$X_2$	$X_4$	$X_5$	$X_3$
	4	$X_2$	$X_2$	$X_1$	$X_5$	$X_4$	$X_3$
	5	$X_1$	$X_5$	$X_3$	$X_4$	$X_2$	$X_1$
	6	$X_4$	$X_2$	$X_4$	$X_5$	$X_3$	$X_1$
	7	$X_2$	$X_5$	$X_3$	$X_2$	$X_1$	$X_4$
	8	$X_2$	$X_5$	$X_2$	$X_1$	$X_3$	$X_4$
	9	$X_3$	$X_2$	$X_1$	$X_4$	$X_3$	$X_5$
	10	$X_1$	$X_5$	$X_4$	$X_1$	$X_3$	$X_2$

Fig. 3. An example of a random pilot schedule for the UE of interest and potential contaminators in a neighboring cell. Green boxes represent pilots, which are orthogonal to the pilot from the UE of interest. Red boxes represent contamination and  $\mathbf{x}_i$  denotes a pilot sequence.

**Example:** To help the understanding of the benefit from pilot sequence hopping, consider the ideal case of a constant channel between BS and UE of interest and a single contaminating neighboring cell. Noise is disregarded in this example, since attention is on decontamination. Moreover, we assume an infinite amount of orthogonal pilot sequences and

an infinite amount of users per cell, such that  $\tau = K = \infty$  and  $\mathbb{E}[t_c] = \infty$ , which means contaminating signals in all time slots are independent. For simplicity, we assume  $\mathbf{x}_n^H \mathbf{x}_n = 1$ , such that the estimate in time slot  $n$  is

$$\hat{h}_n = h + h'_n, \quad (5)$$

where  $h'_n$  is the channel of the contaminator in time slot  $n$ . Now consider a new estimator,  $\bar{h}_n$ , which is the average of all estimates until time slot  $n$ . Hence, we have

$$\bar{h}_n = h + \frac{1}{n} \sum_{i=1}^n h'_i. \quad (6)$$

In this case, the error in the estimate is solely composed of the average of the contaminating signals, which are independent and have variance  $\sigma_c^2$ . Hence, the variance of the estimation error is  $\frac{\sigma_c^2}{n}$ . If pilot sequence hopping had not been performed, the variance of the estimation error had remained  $\sigma_c^2$ , since  $h'_n$  would be constant. Note that the MSE goes towards zero for  $n \rightarrow \infty$ , when pilot sequence hopping is performed. This is a result of the fact that a pilot signal in the infinite past carries as much information about the current channel as the most recent pilot signal, in the ideal example of a constant channel. Note also that for finite  $\tau$  (and  $K$ ) and thereby finite  $\mathbb{E}[t_c]$ , the variance of the estimation error is lower bounded by  $\frac{\sigma_c^2}{K}$ , since only a maximum of  $K$  independent estimates can be achieved. In a more practical example with a time-varying channel, the amount of information carried in a pilot signal decays over time. It is, however, still possible to extract such information using appropriate filtering techniques. For this purpose we have chosen a modified version of the Kalman filter, which is described next.

### B. Modified Kalman Filter

A conventional Kalman filter can be used to track the state,  $\mathbf{b}_n$ , of a system based on observations,  $\mathbf{y}_n$ , where

$$\mathbf{y}_n = \mathbf{C}_n \mathbf{b}_n + \mathbf{d}_n, \quad (7)$$

and  $\mathbf{C}_n$  is the measurement matrix of the system and  $\mathbf{d}_n$  is measurement noise. Moreover, the evolution of the system state must follow

$$\mathbf{b}_n = \mathbf{A}_n \mathbf{b}_{n-1} + \mathbf{v}_n, \quad (8)$$

where  $\mathbf{A}_n$  is the state transition matrix and  $\mathbf{v}_n$  is the process noise. In a conventional application of the Kalman filter,  $\mathbf{A}_n$  is assumed constant and known.

The problem of estimating a time-varying channel based on pilot signals, also termed channel tracking, can be solved using the Kalman filter. The observations as expressed in (2) follow the linear model in (7), where the observation matrix is the transmitted pilot sequence and the tracked state is the

channel coefficient. The evolution of the channel coefficient as expressed by Clarke's model does not follow the model in (8). However, it can be transformed into an autoregressive (AR) model with a finite number of coefficients, which follows the form of (8). If the instantaneous velocity of the user of interest, and thereby the autocorrelation function, are known, the AR coefficients can be found using the Yule-Walker equations [14]. However, this cannot be assumed in our case, hence the AR coefficients must be tracked along with the channel state. For this purpose, we must modify the conventional Kalman filter to include an AR model tracker. A 1<sup>st</sup> order AR model is applied, since experiments tell us this adequately captures the autocorrelation of the system. Therefore, only a single AR coefficient,  $a_n$ , must be tracked.

First we state the conventional Kalman filter [15] in our context, where the AR coefficient is assumed known.

For all  $n$ :

$$\mathbf{e}_n = \mathbf{y}_n - \mathbf{x}_n a_{n-1} \hat{h}_{n-1}, \quad (9)$$

$$\mathbf{R}_n = \mathbf{x}_n p_n \mathbf{x}_n^H + \sigma_n^2 \mathbf{I}_\tau + \sigma_c^2 \mathbf{x}_n \mathbf{x}_n^H, \quad (10)$$

$$\mathbf{k}_n = p_n \mathbf{x}_n^H \mathbf{R}_n^{-1}, \quad (11)$$

$$\hat{h}_n = a_n \hat{h}_{n-1} + \mathbf{k}_n \mathbf{e}_n, \quad (12)$$

$$p_{n+1} = a_n^2 (1 - \mathbf{k}_n \mathbf{x}_n) p_n + (1 - a_n)^2, \quad (13)$$

where  $\sigma_n^2$  and  $\sigma_c^2$  are noise power and total contamination power (average over time), respectively, which are both assumed known,  $\mathbf{I}_\tau$  is the  $\tau \times \tau$  identity matrix and  $\hat{h}_n$  is the estimate of  $h_n$ .

For the tracking of the AR coefficient, an approach similar to the one in [16] is taken. In [16] the inclusion of an AR coefficient tracker is presented for a Kalman predictor, i.e. a filter with the purpose of predicting the channel,  $h_n$ , based on all observations until  $\mathbf{y}_{n-1}$ . In this work, we extend this approach to take all observations until  $\mathbf{y}_n$  into account.

The approach is based on calculating the partial derivative with respect to  $a_n$  of the cost function, the mean squared error (MSE), and using this to adjust  $a_n$  in the direction of decreasing MSE. The partial derivative of the MSE is

$$\begin{aligned} \nabla_n &= \frac{\partial}{\partial a_n} \mathbb{E} [|\mathbf{e}_n|^2] \\ &= -(q_{n-1}^H a_{n-1} \mathbf{x}_n^H + \hat{h}_{n-1}^H \mathbf{x}_n^H) \mathbf{e}_n, \end{aligned} \quad (14)$$

where  $q_n = \frac{\partial \hat{h}_n}{\partial a_n}$  and is found by differentiating (12) with respect to  $a_n$ , such that

$$q_n = (1 - \mathbf{k}_n \mathbf{x}_n) (a_n q_{n-1} + \hat{h}_{n-1}) + \mathbf{m}_n \mathbf{e}_n. \quad (15)$$

Here,  $\mathbf{m}_n = \frac{\partial \mathbf{k}_n}{\partial a_n}$ , which is found by differentiating (11) with respect to  $a_n$ , hence

$$\mathbf{m}_n = (1 - \mathbf{k}_n \mathbf{x}_n) s_n \mathbf{x}_n^H \mathbf{R}_n^{-1}. \quad (16)$$

Finally, we introduced  $s_n = \frac{\partial p_n}{\partial a_n}$ , which is a differentiation of (13) with respect to  $a_n$ , giving us

$$s_{n+1} = a_n^2(1 - \mathbf{k}_n \mathbf{x}_n) s_n (1 - \mathbf{x}_n^H \mathbf{k}_n^H) - 2a_n \mathbf{k}_n \mathbf{x}_n p_n. \quad (17)$$

Using  $\nabla_n$ , we can adjust  $a_n$  as follows

$$a_n = [a_{n-1} - \mu [\nabla_n]_{-\nu}^{+\nu}]_0^1, \quad (18)$$

where  $\mu$  is a parameter adjusting the convergence speed and the brackets denote truncations. The inner truncation involving  $\nu$  is to avoid dramatic adjustments in situations with a high slope and the outer truncation is to obey  $0 \leq a_n \leq 1$ . The need for  $\nu$  will be explained in section IV.

We can now state the modified Kalman filtering algorithm including an AR coefficient tracker:

For all  $n$ :

$$\begin{aligned} \mathbf{e}_n &= \mathbf{y}_n - \mathbf{x}_n a_{n-1} \hat{h}_{n-1}, \\ \mathbf{R}_n &= \mathbf{x}_n p_n \mathbf{x}_n^H + \sigma_n^2 \mathbf{I}_\tau + \sigma_c^2 \mathbf{x}_n \mathbf{x}_n^H, \\ \nabla_n &= -(q_{n-1}^H a_{n-1} \mathbf{x}_n^H + \hat{h}_{n-1}^H \mathbf{x}_n^H) \mathbf{e}_n, \\ a_n &= [a_{n-1} - \mu [\nabla_n]_{-\nu}^{+\nu}]_0^1, \\ \mathbf{k}_n &= p_n \mathbf{x}_n^H \mathbf{R}_n^{-1}, \\ \hat{h}_n &= a_n \hat{h}_{n-1} + \mathbf{k}_n \mathbf{e}_n, \\ \mathbf{m}_n &= (1 - \mathbf{k}_n \mathbf{x}_n) s_n \mathbf{x}_n^H \mathbf{R}_n^{-1}, \\ q_n &= (1 - \mathbf{k}_n \mathbf{x}_n) (a_n q_{n-1} + \hat{h}_{n-1}) + \mathbf{m}_n \mathbf{e}_n, \\ p_{n+1} &= a_n^2 (1 - \mathbf{k}_n \mathbf{x}_n) p_n + (1 - a_n)^2, \\ s_{n+1} &= a_n^2 (1 - \mathbf{k}_n \mathbf{x}_n) s_n (1 - \mathbf{x}_n^H \mathbf{k}_n^H) - 2a_n \mathbf{k}_n \mathbf{x}_n p_n. \end{aligned} \quad (19)$$

#### IV. NUMERICAL RESULTS

The proposed scheme (Estimator) has been simulated and compared to the scheme from [16] (Predictor) and the conventional solutions of least squares (LS) estimation and minimum mean squared error (MMSE) estimation based on a single time slot. The expressions for the LS and MMSE estimators are given in (20) and (21), respectively. An overview of the parameters, which are common for all simulations, is given in Table I. The choice of  $\mu$  is based on experiments showing that this is a good compromise between convergence speed and robustness towards variance. Throughout all simulations, we assume that all users have equal and constant mobility. Moreover, we assume that contaminating signals have zero autocorrelation between time slots, which is justified by the choice of  $K = 96$ , such that  $\mathbb{E}[t_c] = 96$ , cf. (4).

$$\hat{h}_n^{ls} = (\mathbf{x}_n^H \mathbf{x}_n)^{-1} \mathbf{x}_n^H \mathbf{y}_n, \quad (20)$$

$$\hat{h}_n^{mmse} = \mathbf{x}_n^H (\mathbf{x}_n \mathbf{x}_n^H + \sigma_n^2 \mathbf{I}_\tau + \sigma_c^2 \mathbf{x}_n \mathbf{x}_n^H)^{-1} \mathbf{y}_n. \quad (21)$$

Initially, results are shown for the conventional Kalman filter expressed in equations (9) through (13). MSE as a function of the user mobility,  $v$ , and the AR coefficient,  $a_n$ , is shown in Fig. 4. From this figure, it is evident how important it is to have

TABLE I  
SIMULATION PARAMETERS

Parameter	Value	Description
$\sigma_n^2$	0.2	Noise variance
$L$	7	Number of cells
$K$	96	Users per cell
$\tau$	96	Pilot length
$\mu$	$10^{-5}$	Convergence speed
$\nu$	100	Derivative cap
$f_c$	1.8 GHz	Carrier frequency
$N_s$	20	Number of scatterers
$t_s$	0.5 ms	Time between pilots
$a_0$	0.5	Initial AR coefficient
$\hat{h}_0$	0	Initial estimate
$q_0$	0	Initial differentiated estimate
$p_1$	0	Initial error covariance
$s_1$	0	Initial differentiated error covariance

an accurate AR model, which suits the current mobility of the UE of interest. This stresses the need for the modification of the Kalman filter, as proposed in section III-B. Moreover, it is seen that the derivative of the MSE with respect  $a_n$  may attain very high values at low  $a_n$ . This can cause undesirably high variance in the estimate of the optimal  $a_n$ , which motivates the use of a derivative cap,  $\nu$ .

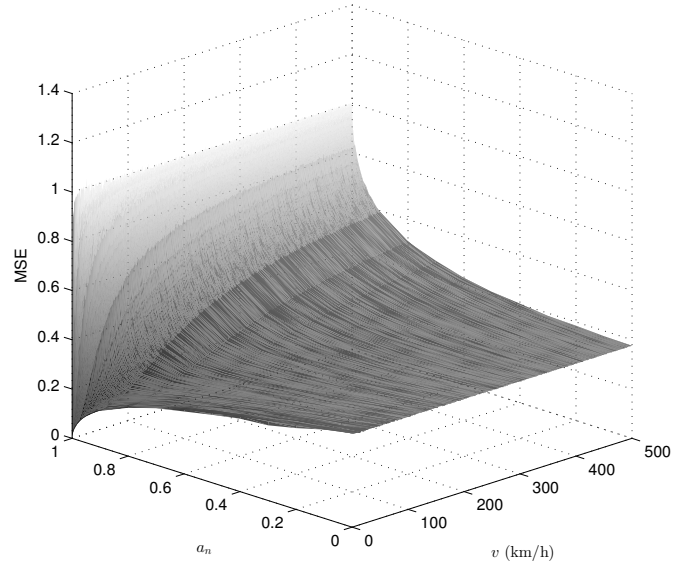


Fig. 4. MSE as a function of the autoregressive model coefficient and the user mobility.

Fig. 5 shows a comparison of the simulated estimators with respect to MSE as a function of user mobility when  $\sigma_c^2 = 0.6$ . For both the predictor and the scheme proposed in this work, results where the optimal value of  $a_n$  is assumed to be known, have been included. This highlights the performance of the tracker. It is evident that the tracker provides a very



good estimate of the optimal AR coefficient. Moreover, it is seen that the proposed scheme outperforms LS and MMSE and performs as well as the predictor at low mobility. At high mobility, the proposed scheme outperforms LS and the predictor, while matching the performance of MMSE.

A different perspective is given in Fig. 6. Here the MSE is plotted as a function of the signal-to-interference ratio (SIR), at typical mobility levels as defined by 3GPP [17]. This figure shows how the proposed scheme is able to suppress even very strong contamination at typical mobility.

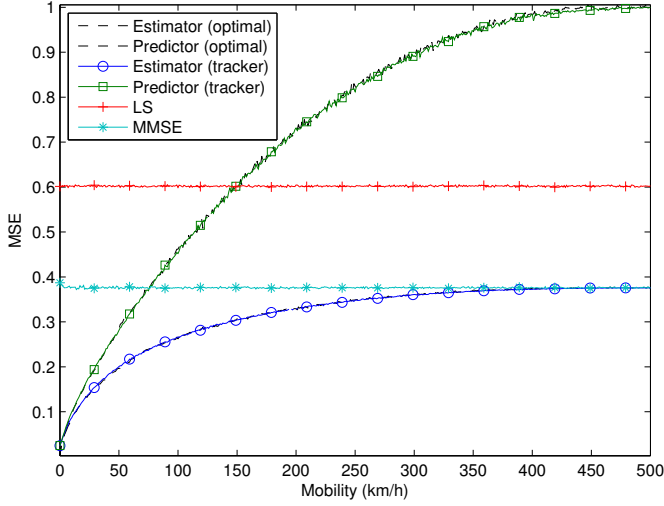


Fig. 5. Comparison between the proposed scheme and conventional solutions with respect to means squared error as a function of mobility.

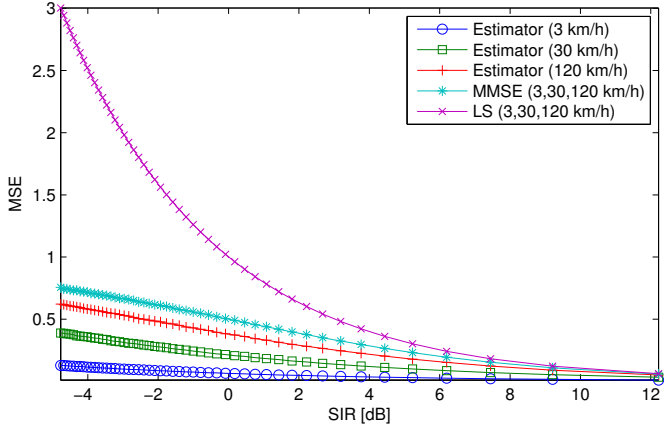


Fig. 6. Comparison between the proposed scheme and conventional solutions with respect to means squared error as a function of the signal-to-interference ratio.

## V. CONCLUSIONS

We have presented a solution to pilot contamination in channel estimation, which is a major challenge in massive MIMO systems. It is based on a combination of a pilot sequence hopping scheme and a modified Kalman filter. The pilot sequence hopping scheme involves random shuffling of

the assigned pilot sequences within a cell, which ensures decorrelation in the time dimension of the contaminating signals. This is essential, since it enables subsequent filtering to suppress the contamination. For this filtering, the Kalman filter has been chosen, due to its ability to track a time-varying state. However, a conventional Kalman filter is not able to adapt to changes in the underlying model, which is necessary when users have unknown and varying levels of mobility. For this problem we have presented a modified Kalman filter, which can adapt the underlying model based on a minimization of the mean squared error.

Numerical evaluations show that the proposed solution can suppress a significant portion of the contamination at low and moderate levels of mobility. Even at high mobility, i.e. car speeds of 100 to 130 km/h, the proposed solution can provide a noticeable gain over conventional estimation methods.

## REFERENCES

- [1] T. Brown, E. De Carvalho, and P. Kyritsi, *Practical Guide to the MIMO Radio Channel*. John Wiley & Sons, 2012.
- [2] T. Marzetta, "How much training is required for multiuser MIMO?," in *Signals, Systems and Computers, 2006. ACSSC '06. Fortieth Asilomar Conference on*, pp. 359–363, Oct 2006.
- [3] F. Rusek, D. Persson, B. K. Lau, E. Larsson, T. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *Signal Processing Magazine, IEEE*, vol. 30, pp. 40–60, Jan 2013.
- [4] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *Selected Areas in Communications, IEEE Journal on*, vol. 31, pp. 264–273, February 2013.
- [5] A. Ashikhmin and T. Marzetta, "Pilot contamination precoding in multi-cell large scale antenna systems," in *Information Theory Proceedings (ISIT), 2012 IEEE International Symposium on*, pp. 1137–1141, July 2012.
- [6] J. Jose, A. Ashikhmin, T. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *Wireless Communications, IEEE Transactions on*, vol. 10, pp. 2640–2651, August 2011.
- [7] H. Q. Ngo and E. Larsson, "EVD-based channel estimation in multicell multiuser MIMO systems with very large antenna arrays," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pp. 3249–3252, March 2012.
- [8] R. Müller, L. Cottatellucci, and M. Vehkaperä, "Blind pilot decontamination," *Selected Topics in Signal Processing, IEEE Journal of*, vol. PP, no. 99, pp. 1–1, 2014.
- [9] R. Müller, M. Vehkaperä, and L. Cottatellucci, "Blind pilot decontamination," in *ITG Workshop on Smart Antennas*, March 2013.
- [10] R. Müller, M. Vehkaperä, and L. Cottatellucci, "Analysis of blind pilot decontamination," in *Proceedings of the 47th Annual Asilomar Conference on Signals, Systems, and Computers*, November 2013.
- [11] L. Cottatellucci, R. Müller, and M. Vehkaperä, "Analysis of pilot decontamination based on power control," in *Vehicular Technology Conference (VTC Spring), 2013 IEEE 77th*, pp. 1–5, June 2013.
- [12] J. Choi, D. Love, and P. Bidigare, "Downlink training techniques for FDD massive MIMO systems: Open-loop and closed-loop training with memory," *Selected Topics in Signal Processing, IEEE Journal of*, vol. PP, no. 99, pp. 1–1, 2014.
- [13] R. Clarke, "A statistical theory of mobile-radio reception," *Bell system technical journal*, vol. 47, no. 6, pp. 957–1000, 1968.
- [14] S. Kay and J. Marple, S.L., "Spectrum analysis - a modern perspective," *Proceedings of the IEEE*, vol. 69, pp. 1380–1419, Nov 1981.
- [15] S. Haykin, *Adaptive Filter Theory*. Pearson Education, Limited, 2013.
- [16] K.-Y. Han, S.-W. Lee, J.-S. Lim, and K.-M. Sung, "Channel estimation for OFDM with fast fading channels by modified Kalman filter," *Consumer Electronics, IEEE Transactions on*, vol. 50, pp. 443–449, May 2004.
- [17] 3GPP, "Spatial channel model for Multiple Input Multiple Output (MIMO) simulations," TR 25.996, 3rd Generation Partnership Project (3GPP), Sept. 2012.